

Data is Forever: Solving a Major Challenge of Data-Driven Government

By: Jenna Sindle, Managing Editor

In 2020, it was estimated that 1.7Mb of data was created every second for every person on earth and this number will only continue to rise. As the data created and stored by federal agencies grows in size, scope, and volume, so too does the need to retain and access that information securely and inexpensively. While multi-cloud environments may have helped mitigate the challenges agencies face with active data, most agencies are still struggling with how to manage data that is no longer in active production, but which still needs to be retained to comply with regulations and, increasingly, to support AI and machine learning.

“As data ages it doesn’t need to be available on

daily basis, but it does need to be retained from anywhere from 25 to 100 years, or even perpetuity if it’s created by the National Archives and Records Administration, the Executive Office of the President, or the Library of Congress,” explained ViON’s Michael Lamb. “The biggest challenge for most agencies is archiving these large data volumes in a way that tiers it effectively and is cost-effective.

Lamb noted that some existing storage solutions – like public cloud storage – might meet the criteria for inexpensive archiving and long-term backup by some definitions but do not always meet the needs of federal agencies. “Public cloud archiving is great,” he explained, “until the moment when



you need to pull data out. The retrieval time can be lengthy and the costs of pulling that data out can be very high.

While most data typically lies untouched after a short period of use, this is changing. “As agencies invest in artificial intelligence (AI) and machine learning (ML) to advance the mission and address complex challenges more quickly, data is being moved in and out of archival mode far more frequently,” said Lamb. “To design robust and accurate algorithms agencies need to work extensively with existing and archived data. Pulling data out of a public cloud archive is just not practical on a financial level for agencies.”

Instead, Lamb encourages agencies to look to the past to solve the challenges of the future. “While agencies have moved data and workloads to the cloud to support interoperability, when data has aged out of daily use archiving to tape is a smart and cost-efficient option,” Lamb said. Not only does a tape archive move data from an expensive storage environment to an inexpensive format, it also makes light work of compliance and regulatory requirements and provides a strategic advantage to agencies in disaster recovery and recovery from ransomware attacks.

In recent years tape storage might have taken a backseat to cloud solutions, but it never really went away Lamb explained. “Tape manufacturers have

continued to invest in the technology and have made significant improvements,” he said. Likewise service offerings have improved as well with tape archive as-a-Service moving the heavy lift of designing, scaling, and operating the backup from agency IT teams to a trusted service partner. “Archive as-a-Service is very similar to the capacity as-a-Service model that agencies are already very familiar with,” Lamb explained. “It allows an agency to start small and build to meet their needs instead of buying outright the capacity they think they’ll need, and then only using a small fraction of that capacity. It’s truly a cost-effective choice. And, as agencies are accessing historical data more frequently to feed AI and ML models, they won’t be hit with egress charges typically associated with a public cloud.

As federal agencies move to adopt data intensive applications like AI that rely on historical data for modeling, the need for cost-effective and reliable access to that data will intensify. Archive as-a-Service is a scalable, affordable solution that ensures access to critical data when and where it is needed and should be part of a complete data management strategy.